OPTIMIZING MPEG-4 CODING PERFORMANCE BY TAKING POST-PROCESSING INTO ACCOUNT

Wan-Fung Cheung and *Yuk-Hee Chan*

Centre for Multimedia Signal Processing
Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

In this paper, a block-based coding scheme is proposed. Unlike those schemes exploited in a conventional codec, the proposed scheme takes the post-processing scheme to be used in the decoder into account during encoding. An example is also provided in this paper to show how this coding scheme jointly optimizes the compression and the blocking-effect elimination when MPEG-4 deblocking filter is exploited in the post-processing stage.

1. INTRODUCTION

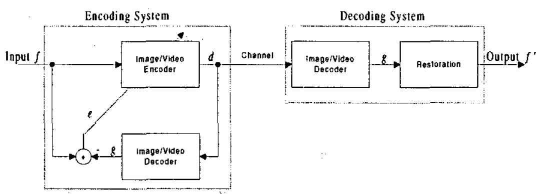


Fig.1 Basic structure of conventional image/video coding system

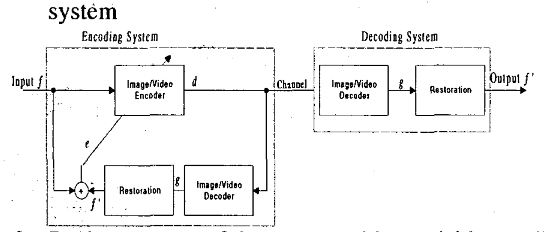


Fig.2 Basic structure of the suggested image/video coding system

In low bit-rate image and video coding, block-based Discrete Cosine Transform (DCT) has been widely used. Noise caused by the coarse quantization of transform coefficients is noticeable as visible discontinuities among adjacent blocks in the reconstructed images. Various post-processing techniques such as block-boundary filtering [1], projections onto convex sets (POCS) [2,3], maximum a posteriori (MAP) methods [4] and transform domain filtering [5,6] have been proposed to remove this blocking artifacts while maintaining the compatibility with current industrial standards. However, in these post-filtering schemes, attention has been focused on the elimination of blocking effect rather than on the optimization of the video codec.

Figure 1 shows the general structure of current image/video encoding systems. In this structure, a post-processing is performed to a decompressed image in the receiver. However, this post-processing stage is not taken into account in the encoder so as to optimize the coding performance. We suggest modifying the basic structure of the encoding system as shown in Figure 2. By doing so, the compression and the post-processing processes can be jointly optimized in the encoder according to f-f'.

In this paper, we will use an example to show how this idea works. We optimize a MPEG-4 [7] codec in a case that a deblocking filter is used in the decoder to eliminate the blocking effect such that the quality of the output will be better than that without optimization.

2. BLOCKING-EFFECT ELIMINATION

In conventional schemes, the MPEG-4 encoder partitions the original image **I** into blocks of size 8x8, say $\mathbf{I}_{m,n}$'s, and then encodes the blocks with DCT block by block with a raster scan strategy. At the decoder side, the deblocking filter [8] performs one-dimensional filtering along the 8x8 block edges to eliminate the discontinuity among adjacent blocks. It is first applied to the horizontal edges and then the vertical edges. If pixel intensity is changed by a previous filtering operation, the updated pixel intensity is used for the next filtering operation. This deblocking filter has two separate filtering modes: DC offset mode and default mode. DC offset mode is applied in smooth region to smooth 8-pixel line segments which are perpendicular to the block's boundary, while default mode is applied in complex regions to smooth 2-pixel line segments.

3. ENCODER OPTIMIZATION

We modify the MPEG-4 encoder according to the post-processing process to combine the compression and the restoration processes to make use of each other to optimize the overall coding performance.

The philosophy of our approach is very simple. Instead of encoding block $\mathbf{I}_{m,n}$ directly, we encode its modified version, say $\mathbf{Y}_{m,n}$, such that $\|F\{(T^{-1}\{Q\{T\{\mathbf{Y}_{m,n}\}\}\})^*\} - \mathbf{I}_{m,n}^*\| < \|F\{(T^{-1}\{Q\{T\{\mathbf{I}_{m,n}\}\}\})^*\} - \mathbf{I}_{m,n}^*\|$. Here, $\mathbf{B}_{m,n}^*$ denotes the greater area of a block named as $\mathbf{B}_{m,n}$, and, operators T{.}, Q{.}, $T^{-1}\{.\}$ and F{.}, respectively, perform the DCT transformation, the quantization, the inverse DCT transformation and the deblocking filtering [8] proposed for MPEG4 standard. The greater area of a block is actually the support region of operation F{.} performed to the block.

The modification of $\mathbf{I}_{m,n}$ is carried out with an optimization filter. Its design corresponds to that of the deblocking filter used in the post-processing stage to adjust the boundary pixels. Accordingly, it is one-dimensional and has two separate operation modes. In particular, it operates in *Opt-smooth filtering mode* to handle smooth region and *Opt-default filtering mode* to handle complex region.

Let $\vec{v} = (v_0, v_1, v_2, \ldots v_8, v_9)$ be a particular pixel vector across the boundary of block $\mathbf{C}_{m,n}$ and block $\mathbf{C}_{m-1,n}$ as shown in Figure 3, where $\mathbf{C}_{m,n}$ is the MPEG-coded version of $\mathbf{Y}_{m,n}$. At the decoder, this pixel vector will be post-processed with the one-dimensional deblocking filter suggested in [8].

The mean square deviation of the post-processing result, say $\vec{p}$, from the corresponding pixel vector in the original image, say $\vec{i}$, is then given by $\xi = \| (\vec{p} - \vec{i}) \|^2$. The values of $\vec{v}_2$ which minimize ξ are determined by solving equations $\partial\xi / \partial\vec{v}_2 = 0$. Note only pixels in the current processing block are used to minimize ξ. Pixel vectors in the processing block are processed accordingly one by one in a predefined order until the whole desirable $\mathbf{C}_{m,n}$ is obtained. If a pixel was modified, its most updated value will be used as the input for the current processing.

After obtaining the desirable $\mathbf{C}_{m,n}$, $\mathbf{Y}_{m,n}$ can then be determined. In theory, $\mathbf{Y}_{m,n}$ should be an array that makes $\mathbf{C}_{m,n} = T^{-1}\{Q\{T\{\mathbf{Y}_{m,n}\}\}\}$. However, this cannot be practically achieved as quantization is involved and $\mathbf{C}_{m,n}$ may not be one of the possible DCT-coded outputs provided by the encoder. In our approach, we let $\mathbf{Y}_{m,n} = \mathbf{C}_{m,n}$.

Sometimes the proposed modification may not guarantee that $F\{(T^{-1}\{Q\{T\{\mathbf{Y}_{m,n}\}\}\})^*\}$ is more faithful to $\mathbf{I}_{m,n}^*$ than $F\{(T^{-1}\{Q\{T\{\mathbf{I}_{m,n}\}\}\})^*\}$. Hence, the encoder should determine whether $\mathbf{Y}_{m,n}$ or $\mathbf{I}_{m,n}$ should be encoded. The selection is carried out by checking their post-processed results against the original image. The one provides the minimum mean square error (MSE) and less bit stream is then encoded and transmitted. We proceed to process and encode the next block until the whole image is encoded.

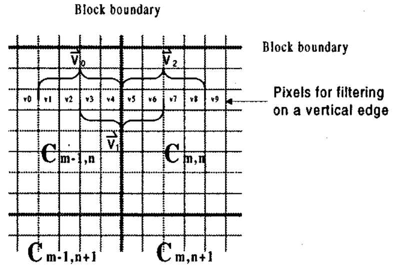
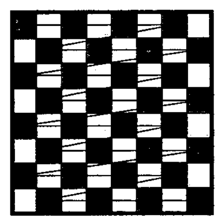


Fig. 3 Block boundaries of 8x8 blocks

4. OPTIMIZATION ORDER

In order to minimise the interference among the modification results, an *Optimization filtering with Checker scheme (OPF-CK)* is used to modify $\mathbf{I}_{m,n}$ to $\mathbf{Y}_{m,n}$. In this scheme, blocks are processed in an order as shown in Figure 4. We process and encode the blocks at the gray color area first and then white color area until the whole image is encoded. The optimization filtering starts at the left-lower corner of a block and moves clockwise around the block. Hence, we can modify all pixel vectors around the boundaries of the block. Besides, the blocks in the area of the same colour can be processed in parallel to reduce processing time as the modification of these blocks does not affect each other.

As the modification made in a particular block will influence the deblocking result of some previously modified blocks next to it, the pre-processing is performed in an iterative way in order to reach an optimized output. From our simulation result, we found that it converged rapidly and was able to reach a nearly optimized rate-distortion value after 2 iterations. All results shown in the paper are obtained with 2 iterations unless it is specified.



(Arrow shows the scanning path)

Fig.4 Checker scans strategy

5. SIMULATION RESULT

Computer simulation has been carried out to demonstrate that a corresponding pre-processing step can definitely improve the coding performance if we take the post-processing step into account. In our simulation, the MPEG-4 video VM 9.1 coder [7] is used. We consider I-VOP and P-VOP with a fixed quantization parameter (QP). The ITU-T Recommendation H.263 [9] quantization method is adopted and the motion search range is [-16., 15.5]. Various QCIF test sequences of size 150 frames were used to evaluate the performance of the proposed scheme. Only the first frame of a sequence is coded as intra-frame.

Table 1 shows the simulation results of 5 testing video sequences. One can see that the PSNR and the compression rate can be further improved if a pre-processing step is performed prior to encoding and the proposed pre-filtering scheme is robust with respect to various video sequence characteristics.

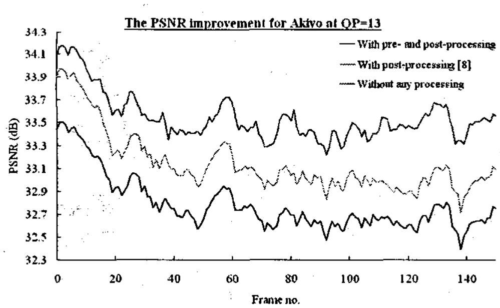
Figure 5 shows the PSNR improvement for *Akiyo* and *Silent* sequences at a fixed QP. Though the improvement seems little in terms of absolute figure, the pre-processing step actually introduces a twofold increase in the improvement in general. Similar results were obtained with different testing video sequences and different sets of QP parameters. Figure 6 shows the rate-distortion performance of different approaches. One can see that the improvement is very respectable at low bit rates. Figure 7 shows the deblocking results for the *Bream* sequence. The proposed pre-processing scheme can preserve the details while reducing blocking artifacts. This can be observed by examining the texture detail around the anal fin and the tail of the fish in Figure 7.

6. CONCLUSION

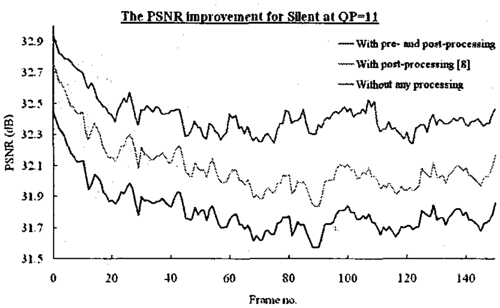
In this paper, a pre-processing method is proposed to improve the coding performance of MPEG-4 by jointly optimizing both compression and blocking-effect elimination. The simulation results show that the proposed method can improve image quality subjectively and objectively without changing the core structure and the bit-stream format of a MPEG4 codec.

7. ACKNOWLEDGEMNTS

The work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Project No. PolyU 5092/98E)

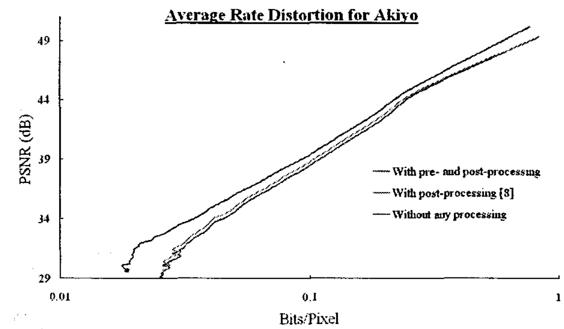


(a)

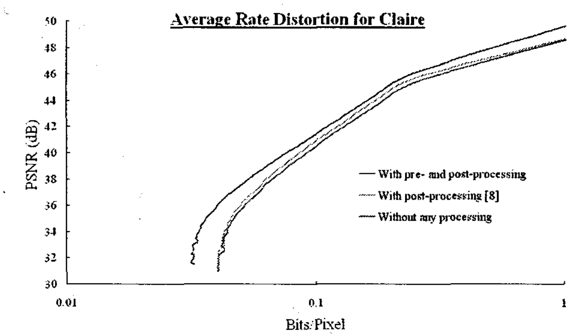


(b)

Fig. 5 PSNR improvement of different coding schemes: (a) *Akiyo* sequence at QP=13 and (b) *Silent* sequence at QP=11.



(a)



(b)

Fig. 6 Rate-distortion performance of different coding schemes: (a) A*kiyo* and (b) *Claire*.

Sequenc	QP	Average PSNR [bits] per frame		
		Without post-processing	With post-processing	
			With Pre-processing	Without Pre-processing
Akiyo	13	32.78dB [1353]	33.14dB [1353]	33.54dB [101
Bream	16	31.34dB [3272]	31.49dB [3272]	31.76dB [3142]
CarPhon	12	32.53dB [3583]	32.80dB [3583]	33.13dB [3533]
Foreman	17	29.80dB [3951]	29.96dB [3951]	30.17dB [3888]
Silent	10	32.34dB [2958]	32.59dB [2958]	32.94dB [2778]

Table 1 Rate-distortion performance of various coding schemes.

8. REFERENCE

[1] Y.H. Chan, S.W. Hong and W.C. Siu, "A practical post-processing technique for real-time block-based coding system," *IEEE trans. on CAS for Video Technology*, Vol.8, No.1, Feb 1998, pp.4-8.

[2] Y.Yang, N.P.Galatsanos, and A.K. Katsaggelos, "Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images", *IEEE Trans. CAS Video Technol.*, vol. 3, no. 6, Dec 1993, pp. 421-432.

[3] Y. Yang, N.P. Galatsanos, and A.K. Katsaggelos, "Projection-based spatially adaptive reconstruction of block-transform compressed images", *IEEE Trans. CAS Video Technol.*, vol. 4, no. 7, Jul 1995, pp. 896-908.

[4] T.P. O'Rourke and R.L. Stevenson, "Improved image decompression for reduced transform coding artifacts" *IEEE Trans. CAS. Video Technol.*, vol. 5, Dec 1995, pp. 490-499.

[5] S.S.O. Choy, Y.H. Chan and W.C. Siu, "Reduction of block-transform image coding artifacts by using local statistics of transform coefficients", *IEEE Signal Processing Letters*, Vol.4, No. 1, Jan 1997, p.5-7.

[6] Z. Xiong, M.T. Orchard, and Y.Q. Zhang, "A deblocking algorithm for JPEG compressed images using overcomplete wavelet representations", *IEEE Trans. CAS. Video Technol.*, vol. 7, no. 2, Apr 1997, pp. 433-437.

[7] "MPEG-4 video verification model V.9.1," ISO/IEC JTC1/SC 29/WG11/N2502a, Oct 1998.

[8] S.D. Kim, J. Yi, H.M. Kim and J.B. Ra, "A deblocking filter with two separate modes in block-based video coding", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, Feb 1998, pp. 156-160.

[9] ITU Telecom. Standardization Sector of ITU, "Video coding for low bit-rate communication," *ITU-T Recommendation H.263*, Mar 1996.

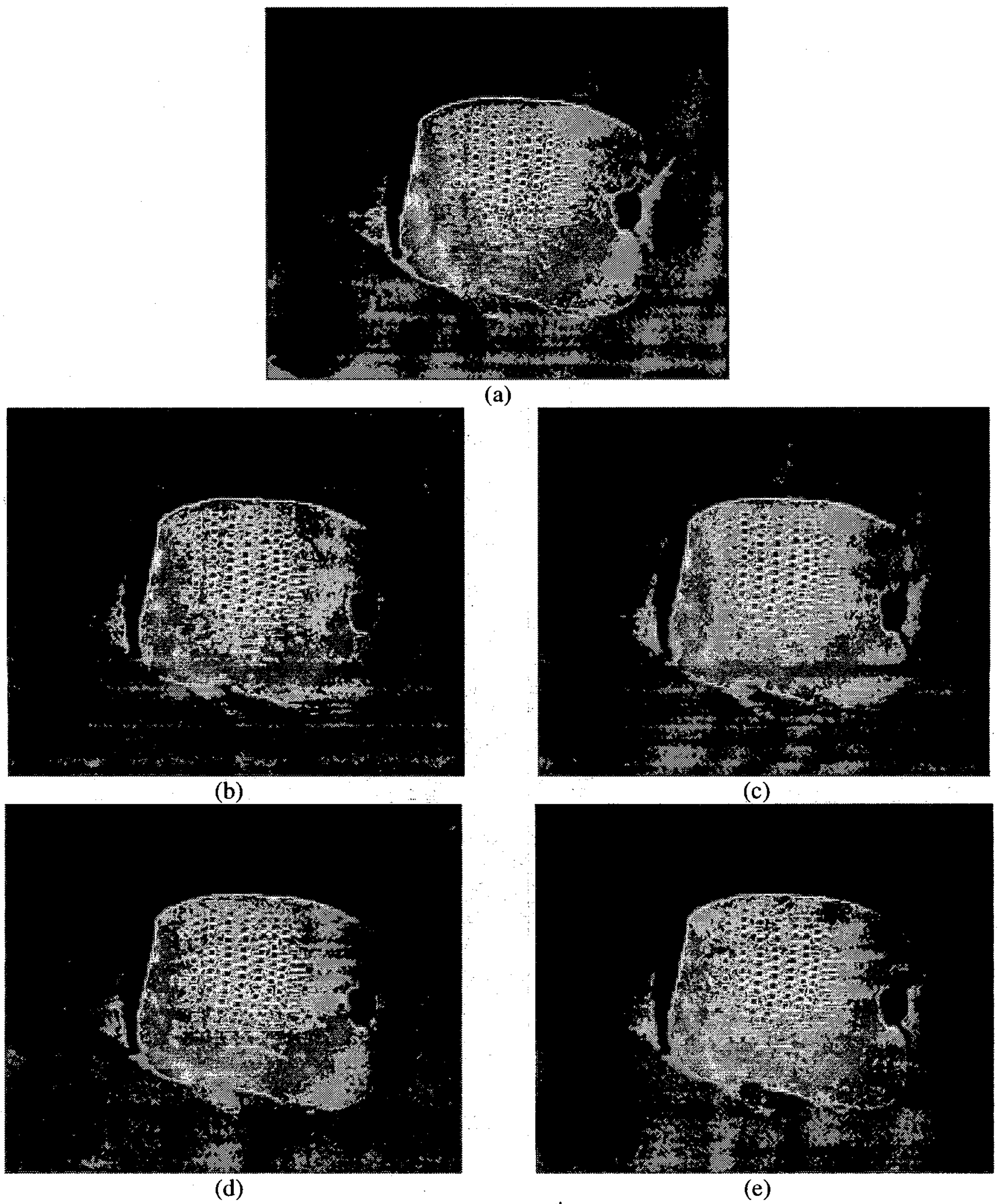


Fig. 7 Coding results for *Bream* sequence (QP=12): (a) the 140th original frame; (b) encoded image of (a) (without pre-processing); (c) post-processing result of (b); (d) encoded image of (a) (with pre-processing); (e) post-processing result of (d).